



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

10821 U.S. PTO

09/855627



FR 600048
#6
45

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

00401368.6

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE,
LA HAYE, LE

07/03/01

This Page Blank (uspto)



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.:
Demande n°: 00401368.6

Anmeldetag:
Date of filing: 18/05/00
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
Koninklijke Philips Electronics N.V.
5621 BA Eindhoven
NETHERLANDS

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:

MPEG-4 binary shape transmission

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:

/

Am Anmeldetag benannte Vertragsstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE/TR
Etats contractants désignés lors du dépôt:

Bemerkungen:
Remarks:
Remarques:

This Page Blank (uspto)

MPEG-4 binary shape transmission**FIELD OF THE INVENTION**

5

The present invention relates to a method of processing a digital video data signal containing data related to rectangular pictures, said method of processing comprising a segmentation step of the digital video data signal for providing segmented video data signals, a segmented video data signal containing a video object which is a region of the rectangular picture. The present invention also relates to a device corresponding to said processing method.

10

Such a method of processing may be used, for example, for encoding a digital video data signal using a video object based encoding framework, such as MPEG-4 encoding standard.

15

BACKGROUND OF THE INVENTION

A video object based encoding framework, such as MPEG-4 encoding standard, referred to as MPEG-4 Visual Version 1, ISO/IEC 14496-2, allows to encode video objects having various shapes instead of the whole rectangular picture. Rectangular pictures are represented by pixels having luminance and chrominance values. In addition to these values, a pixel of a video object has a binary shape value. This value is obtained from a rectangular picture by a segmentation process and is represented by one bit indicating if the pixel is in the object or not. The separate encoding of the video objects may enrich the user interaction in several multimedia services due to flexible access to the digital video data signal and an easy manipulation of the video information. In this framework, the encoder may perform a locally defined pre-processing aimed at the automatic identification of the objects appearing in a sequence of pictures.

20

25

30

35

The operation of segmentation aimed at partitioning a rectangular picture or a video sequence of pictures into regions extracted according to a given criterion. Fig. 1 shows an example of a segmentation process in which a rectangular picture (RP) has been partitioned in several video objects (VO1 to VO4). In the case of a video sequence, this partition should achieve the temporal coherence of the resulting sequence of objects masks representing the video object. Different methods have been proposed for segmentation of video sequences, based on either a spatial homogeneity, a motion coherence criterion or a spatio-temporal processing. These algorithms are expected to identify classes of moving objects according to the luminance homogeneity and the motion coherence criterion.

SUMMARY OF THE INVENTION

It is an object of the invention to provide a method of processing a digital video data signal for providing a modified digital video data signal containing binary shape data.

For the moment, only pixel data transmission is standardised by the recommendation ITU-R BT.601-5. This recommendation specifies methods for digitally encoding video signal but does not propose or suggest any transmission method of the binary shape data.

The method of processing in accordance with the invention is characterised in that it comprises an identification step by an identifier, from the segmented video data signals, to which video object a pixel of the rectangular picture belongs, and an insertion step of the identifiers within the digital video data signal, forming a modified digital video data signal intended to be encoded by a video object based encoding framework.

Such a method of processing allows to insert information related to binary shape data in a digital video data signal by means of identifiers of video objects. As a consequence, the modified digital video data signal obtained by such a method of processing can be directly encoded by a video object based encoder and more especially a hardware encoder.

In the preferred embodiment of the invention, the digital video data signal is defined by the recommendation ITU-R BT.601-5 and the identifiers are first inserted within an ancillary data packet as defined in the recommendation ITU-R BT.1364, which is then inserted within a vertical blanking space of the digital video data signal at a row level.

The present invention also applies to a processing device for implementing such a method of processing.

These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described, by way of example, with reference to the accompanying drawings, wherein :

- Fig. 1 shows an example of a segmented picture comprising various video objects,
- Fig. 2 is a block diagram of a method of processing in accordance with the invention,
- Fig. 3 represents a digital video data signal as defined by the recommendation ITU-R BT.601-5, and

- Fig. 4 represents an ancillary data packet as defined in the recommendation ITU-R BT.1364.

DETAILED DESCRIPTION OF THE INVENTION

5

The present invention aims at inserting binary shape data in a digital video data signal, the modified digital video data signal thus obtained being directly encoded by a video object based encoder. Fig. 2 is a block diagram giving the principle of a method of processing in accordance with the invention.

10

Such a method of processing processes a digital video data signal (DVS) containing data related to rectangular pictures, and segmented video data signals (SVS) provided by a segmentation step (SEG) of the digital video data signal, a segmented video data signal containing a video object (VO) which is a region of the rectangular picture.

Said method of processing comprises the following steps of :

15

- identifying (ID) by an identifier, from the segmented video data signals (SVS), to which video object a pixel of the rectangular picture belongs,
- inserting (INS) the identifiers in the digital video data signal, forming a modified digital video data signal (DVS_m), and
- encoding (ENC) the modified digital video data signal using the MPEG-4

20

encoding standard for providing an encoded data signal (ES).

25

In the preferred embodiment of the invention, the digital video data signal (DVS) is the one defined by the recommendation ITU-R BT.601-5. Fig. 3 shows the structure of a digital video data signal as defined by said recommendation. Such a digital video data signal comprises :

- video data ($YC_R C_B[1]$ and $YC_R C_B[2]$), comprising luminance samples (Y) and two simultaneous colour-difference signals (C_R and C_B),
- horizontal blanking spaces (HBSu1, HBSd1, HBSu2 and HBSd2),
- vertical blanking spaces (VBS1 and VBS2).

30

For example, in a 50 fields per second system, where the whole picture comprises 625 lines, the video data are divided into two fields comprising respectively 288 lines. The rest of the lines corresponds to the various horizontal blanking spaces.

35

If the sampling frequency is 13.5 MHz for the luminance signal, the sampling frequency is 6.75 MHz for each colour difference signal in the 4:2:2 encoding format. The number of samples per total line is 864 for the luminance signal and 432 for each colour-difference signal. These samples are encoded on 8 bits (optionally 10). As the number of

samples per digital active line is 720 for the luminance signal and 360 for each colour-difference signal, 288 samples are at the maximum available for the vertical blanking spaces.

The present invention is applicable for other formats of the digital video data signal as defined by the recommendation ITU-R BT.601-5 such as, for example, a 60 fields per second rate corresponding to a 525-line system, a 4:4:4 encoding format or a sampling frequency of 18 MHz for the luminance signal.

The present invention stays also applicable for other digital video data signals, such as, for example, the ones defined by the recommendation ITU-R BT.656, ITU-R BT.799 or ITU-R BT.1120 corresponding to HDTV signals.

For the purpose of being processed by the processing method, the digital video data signal (DVS) has to be previously segmented using a segmentation process (SEG), and resulting in several segmented video data signals (SVS). The segmentation process can be performed in two ways. The first one is based on an usual software method such as the one depicted in the background of the invention but it takes quite a lot of time. The second one is much faster and is called the Chroma Key process. Such a process is dedicated to the extraction of at least two video objects from which one is the background video. This background is preferably blue or green and such a segmentation process can be implemented in a hardware application.

The identifiers of video objects are then inserted within the digital video data signal using ancillary data as defined in the recommendation ITU-R BT.1364. The ancillary data are carried in packets, each packet carrying its own identification. Fig. 4 shows an ancillary data packet as defined in the recommendation ITU-R BT.1364. Said ancillary data packet comprises :

- an ancillary data flag (ADF) which is a fixed preamble that enables an ancillary data packet to be detected,
- a data identification word (DID) to enable packets carrying a particular type of ancillary data to be identified,
- a data block number (DBN) which is incremented by one for each consecutive data packet sharing a common data identification word and requiring continuity indication,
- a data count word (DC) to indicate the packet length,
- a user data word (UDW) which contains the ancillary data , up to 255 words in each packet,
- a checksum word (CD) used to determine the validity of the ancillary data packet from the data identification word through the user data word.

The recommendation ITU-R BT.1364 provides a mechanism for the transport of ancillary data signals through digital video component interfaces in the digital blanking portion of the digital video data signal. In the preferred embodiment of the invention, ancillary data packets are inserted within vertical blanking spaces (VBS1 and VBS2) of the digital video data signal (DVS) at a row level. Sufficient space is available for the entire packet to be inserted within the same vertical blanking space.

Every pixel row or line represents 720 pixels and the size of the user data word shall not exceed 255 words or bytes. As a consequence, up to 4 video objects (VO) can be inserted in the digital video data signal (DVS). To this end, the method of processing in accordance with the invention comprises an identification step (ID) by an identifier, from the segmented video data signals, to which video object a pixel of the rectangular picture belongs. The video objects are encoded with an identifier having 2 bits. Therefore, 1440 bits, corresponding to 180 bytes, are necessary to fully describe a pixel row.

Said identifier allows to determine to which video object the corresponding pixel belongs as follows :

- 00 : the pixel belongs to the first video object (VO1),
- 01 : the pixel belongs to the second video object (VO2),
- 10 : the pixel belongs to the third video object (VO3),
- 11 : the pixel belongs to the fourth video object (VO4).

The bytes of the user data word are numbered from 0 to 179. The eight bits of the byte numbered n contains the following information :

- the bits 0 and 1 contains the identifier of the pixel $4n$,
- the bits 2 and 3 contains the identifier of the pixel $4n+1$,
- the bits 4 and 5 contains the identifier of the pixel $4n+2$,
- the bits 6 and 7 contains the identifier of the pixel $4n+3$.

Finally, the sub-step of inserting (ADP) the identifiers within an ancillary data packet combined with the sub-step of inserting (VBS) the ancillary data packet within a vertical blanking space, allows to form a modified digital video data signal (DVS_m) intended to be directly encoded by a video object based encoder.

It will be obvious that the verb "comprise" does not exclude the presence of other steps or elements besides those listed in any claim.

CLAIMS

1. A method of processing a digital video data signal (DVS) containing data related to rectangular pictures, said method of processing comprising a segmentation step
5 (SEG) of the digital video data signal for providing segmented video data signals (SVS), a segmented video data signal containing a video object (VO) which is a region of the rectangular picture, **characterised in that** said method of processing comprises an identification step (ID) by an identifier, from the segmented video data signals, to which video object a pixel of the rectangular picture belongs, and an
10 insertion step (INS) of the identifiers within the digital video data signal, forming a modified digital video data signal (DVSm) intended to be encoded by a video object based encoding framework.
2. A method of processing a digital video data signal (DVS) as claimed in claim 1,
15 **characterised in that** the digital video data signal is defined by the recommendation ITU-R BT.601-5 and the insertion step (INS) comprises a first sub-step of inserting (ADP) the identifiers within an ancillary data packet as defined in the recommendation ITU-R BT.1364, and a second sub-step of inserting (VBS) the ancillary data packet within a vertical blanking space of the digital video data signal at a row level.
- 20 3. A method of processing a digital video data signal (DVS) as claimed in claim 1, **characterised in that** the identification step (ID) is intended to give an identifier coded on two bits to a given pixel of the rectangular picture.
4. A device for processing a digital video data signal (DVS) containing data related to rectangular pictures, said processing device comprising means for segmenting (SEG)
25 the digital video data signal to provide segmented video data signals (SVS), a segmented video data signal containing a video object (VO) which is a region of the rectangular picture, **characterised in that** said processing device comprises means for identifying (ID) by an identifier, from the segmented video data signals, to which video object a pixel of the rectangular picture belongs, and means for inserting
30 (INS) the identifiers within the digital video data signal, forming a modified digital video data signal (DVSm) intended to be encoded by a video object based encoding framework.
5. A processing device as claimed in claim 4, **characterised in that** the digital video data signal is defined by the recommendation ITU-R BT.601-5 and the inserting
35 means (INS) are intended to first insert the identifiers within an ancillary data packet (ADP) as defined in the recommendation ITU-R BT.1364, which is then inserted within a vertical blanking space (VBS) of the digital video data signal at a row level.

6. A processing device as claimed in claim 4, **characterised in that** the identifying means (ID) are intended to give an identifier coded on two bits to a given pixel of the rectangular picture.
- 5 7. A digital video data signal as defined by the recommendation ITU-R BT.601-5 comprising ancillary data packets as defined in the recommendation ITU-R BT.1364, an ancillary data packers being inserted within a vertical blanking space of the digital video data signal at a row level **characterised in that** the ancillary data packet comprises identifiers corresponding to video objects, said video objects resulting from a segmentation process of rectangular pictures contained in the digital video
- 10 data signal.

MPEG-4 binary shape transmission**ABSTRACT**

The present invention relates to a method of processing a digital video data signal (DVS) aiming at inserting binary shape data in the digital video data signal. Such a method of processing processes the digital video data signal containing data related to rectangular pictures, and segmented video data signals (SVS) provided by a segmentation step (SEG) of the digital video data signal, a segmented video data signal containing a video object (VO) which is a region of the rectangular picture. Said method of processing comprises the steps of identifying (ID) by an identifier, from the segmented video data signals (SVS), to which video object a pixel of the rectangular picture belongs, inserting (INS) the identifiers in the digital video data signal, forming a modified digital video data signal (DVSm), and encoding (ENC) the modified digital video data signal using a video object based encoding framework for providing an encoded data signal (ES).

Use: MPEG-4 encoding

Reference: Fig.2

1/2

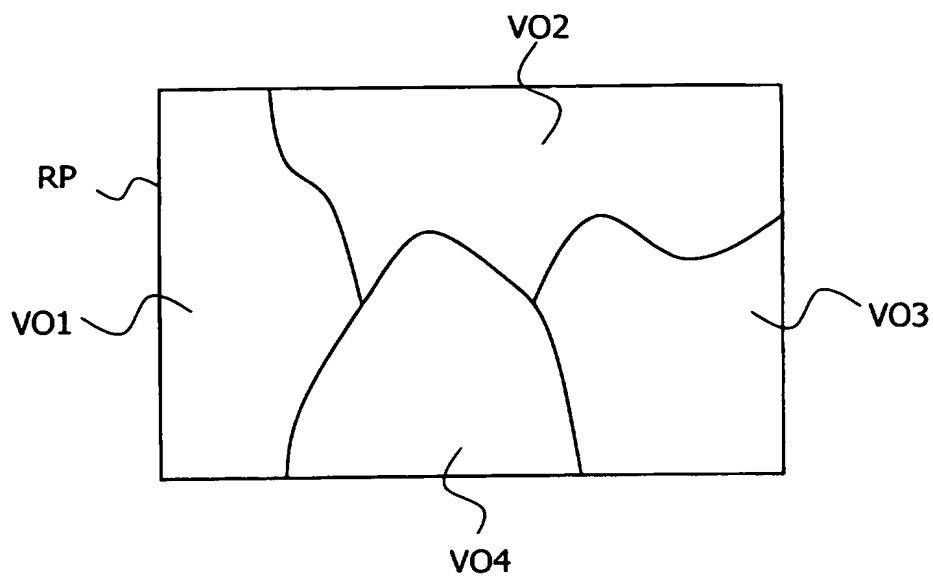


FIG. 1

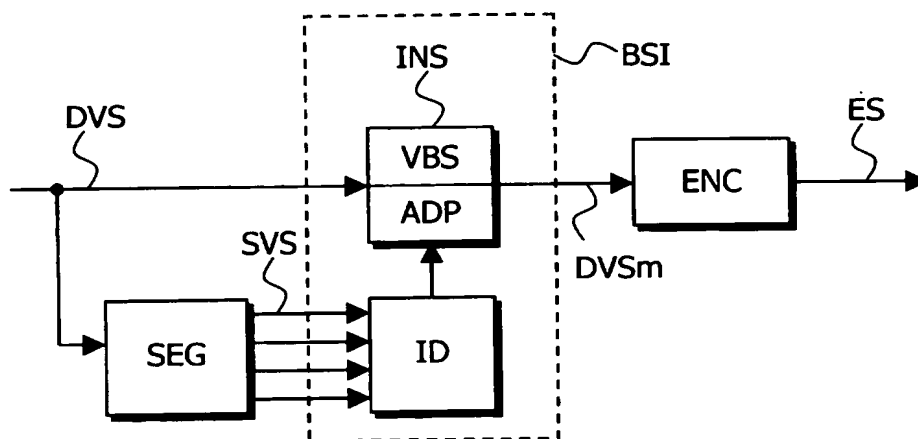


FIG. 2

2/2

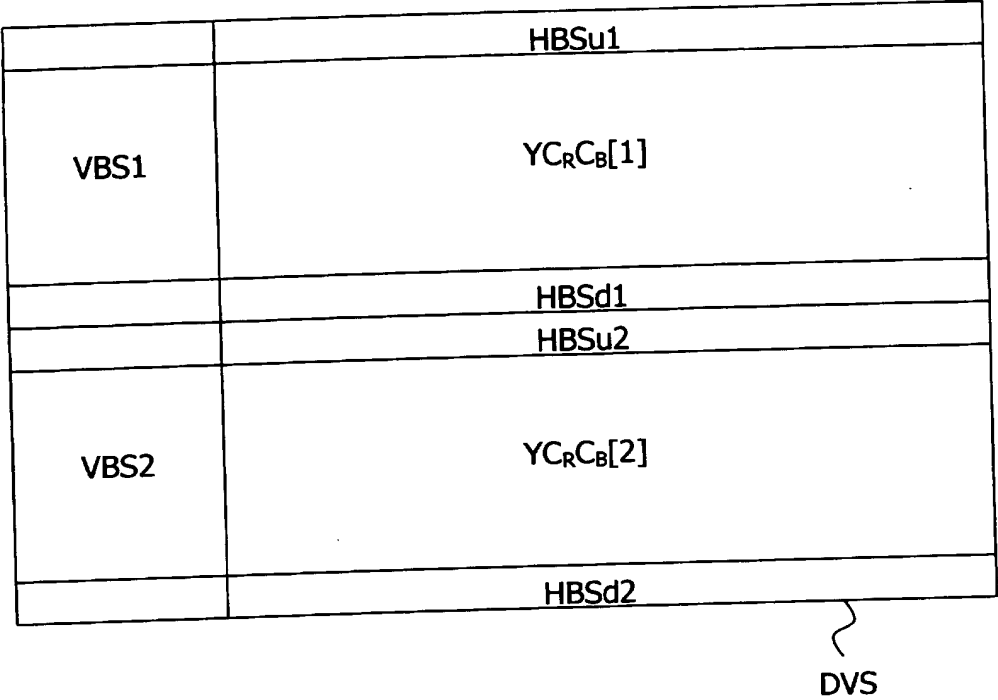


FIG. 3

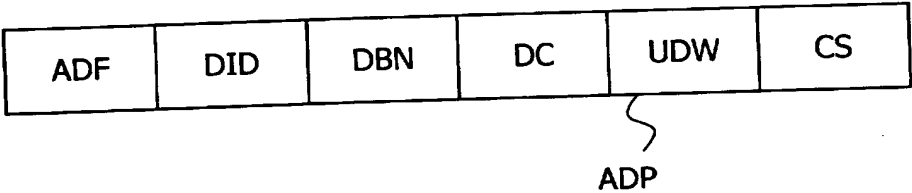


FIG. 4